DISCUSSION PAPER SERIES

No. 8445

COVERAGE OF INFERTILITY TREATMENT AND FERTILITY OUTCOMES: DO WOMEN CATCH UP?

Matilde Pinto Machado
and Anna Sanz-de-Galdeano

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP8445.asp**

ISSN 0265-8003

COVERAGE OF INFERTILITY TREATMENT AND FERTILITY OUTCOMES: DO WOMEN CATCH UP?

Matilde Pinto Machado, Universidad Carlos III de Madrid and CEPR
Anna Sanz-de-Galdeano, Universitat Autònoma de Barcelona

Discussion Paper No. 8445
June 2011

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8445

June 2011

ABSTRACT

Coverage of Infertility Treatment and Fertility Outcomes: Do Women Catch Up?*

The ageing of first-time mothers and the changes in women's labor market conditions have been accompanied by the introduction and subsequent increase in the use of assisted reproductive therapies (ART) that help extend women's reproductive life. Considering the financial cost of infertility treatments, policy interventions that increase insurance coverage may significantly affect fertility trends, and ultimately, population age structures. However, policies have ignored the overall impact of ART coverage on fertility. In this paper, long-term effects of insurance coverage for infertility on the timing of first births and on total fertility rates are examined. Variation in the enactment of infertility insurance mandates over time and across U.S. states allows the estimation of both the short-term and long-term effects. We concentrate on the effects of the more demanding mandates enacted in six states in the later 80s and 90s. Our results show that the effect of these mandates to cover infertility treatment is positive on the average age at first birth and increases over time. The long-term estimates of the increase in age of first-time mothers range from 3 to 5 months. Importantly, we also show that these mandates do not increase the total fertility rates of women by the end of their reproductive lives.

JEL Classification: I18 and J13
Keywords: assisted reproductive technologies (ART), effects of insurance coverage for infertility, insurance coverage for infertility, insurance mandates and total fertility

Matilde Pinto Machado
Department of Economics
Universidad Carlos III de Madrid
Calle Madrid 126
28903 Getafe (Madrid)
SPAIN

Email:
matilde.machado@uc3m.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=150060

Anna Sanz-de-Galdeano
Barcelona GSE
Universitat Autònoma de Barcelona
Office B3-194, Building B
Campus de Bellaterra
08193 Bellaterra, Barcelona
SPAIN

Email:
anna.sanzdegaldeano@uab.cat

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=163119

* We are grateful to Alberto Abadie, Manuel Arellano, Manuel Bag¨ues, Marianne Bitler, Giorgio Brunello, Guillermo Caruana, Monica Costa-Dias, Joao Ejarque, Roger Feldman, Eugenio Giolito, Libertad González, Rachel Griffith, Andrea Ichino, Julián Messina, Pedro Mira, Enrico Moretti, Nuno Sousa Pereira, Helena Szrek, Marcos Vera and Ernesto Villanueva for helpful comments as well as the audiences at the 4th COSME Workshop on Gender Economics (2011), ASHEcon (2010), Health Econometrics Workshop (2010), ESPE (2009), Barcelona GSE Trobada (2009), 3rd PEJ conference, EARIE (2009), SAE (2009), and APES (2009). Matilde P. Machado acknowledges financial support from the Spanish Ministry of Science and Technology Grant SEJ2007-66268. Anna Sanz-de-Galdeano is also an IZA Research Fellow, a MOVE Research Fellow and is affiliated with the Barcelona Graduate School of Economics (Barcelona GSE). She acknowledges financial support from the Spanish Ministry of Science and Technology Grant SEJ2007-62500, the Government of Catalonia (Contract no. 2009SGR189, the XREPP and the Barcelona GSE Research Network).

Submitted 10 June 2011

1 Introduction

The average age at first birth in the United States has been rising steadily over the past decades, from 21.49 in 1968 to 23.72 in 1985 and 25.26 in 2004. As shown in Fig. 1, this increase has been accompanied by remarkable changes in the age distribution of first-time mothers, which has become less skewed with a substantially higher density after age 25 and an extension of first-time motherhood beyond age 40.

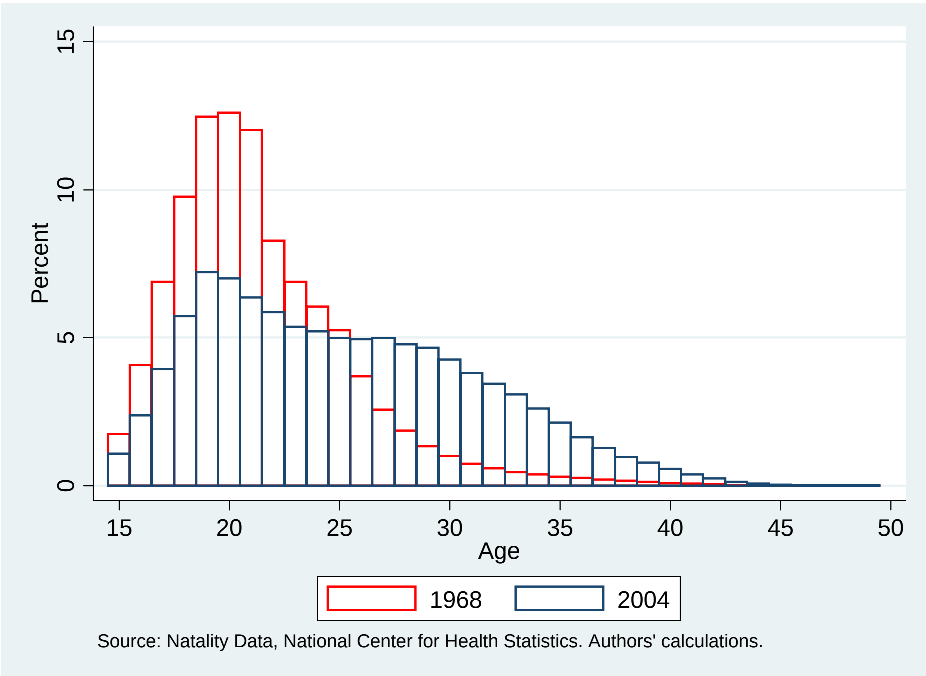

Figure 1: Distributions of maternal age at first birth in 1968 and 2004

Fertility postponement has been studied in conjunction with the increase in women's labor force participation and the changes in their labor market prospects (Blau and Kahn, 1997; Olivetti, 2006); the widespread use of oral contraceptives and its consequences on women's careers (Goldin and Katz, 2002); and an increase in the returns to women's labor market experience (Caucutt et al., 2002).

Women, however, face a biological time constraint on bearing children because fecundity decreases with age. Hence, the ageing of first-time mothers and the changes in women's labor market conditions have been accompanied by the introduction and subsequent increase in the use of Assisted Reproductive Therapies (ART) that help extend women's reproductive life (CDC, 2007). ART techniques, such as *in-vitro* fertilization (IVF), have been available since the late 1970's. The first successful IVF procedure in the United States was achieved in 1981. ART techniques, particularly IVF, are very expensive procedures. For example, in 1992, an IVF delivery cost between $44,000$ and $211,942$ USD (Neumann *et al.*, 1994). However, the costs of ART have decreased substantially in recent decades for various reasons. First, the growing number of infertility clinics throughout the country has resulted in lower prices (Hamilton and McManus, 2005) and in shorter distances that prospective parents need to travel. Second, technological advances have decreased the number of cycles[1] needed per live-birth.[2] Third, insurance coverage for infertility treatments has grown both in the US and in Europe.[3] By 2001, the use of ART had increased so that more than 1% of live births in the U.S. were due to IVF (CDC, 2007).

This paper answers two questions crucial for understanding the overall impact of ART insurance coverage. First, does the coverage of ART have long-term effects on the average age of first-time mothers? Second, does the coverage of ART increase total fertility by the end of a woman's reproductive life? (*i.e.,* do women with easier access to ART end up having more children than their counterparts with less access due to the higher prevalence of multiple births, for example)?[4]

Considering the high cost of infertility treatments (Bitler and Schmidt, 2011; Collins, 2001), policy interventions that grant insurance coverage for infertility treatments may affect fertility trends, and ultimately, population age structures. This study contributes to ongoing debates about infertility treatments in the U.S. as well as in Europe, two regions with very different health systems. Results for the mid to long-term consequences of ART are central to the European debate on possible solutions to an ageing population, *i.e.*, can ART be part of a package of policies intended to increase fertility rates in Europe? (Grant, 2006, Ziebe and Devroey, 2008).[5] The answer to this question is complex because the short-term effect of an increase in the coverage for infertility treatment may be very different from the long-term effect.

In the short-term, an increase in the aggregate fertility rate is usually expected due to an increase

[1] A *cycle* is the process that starts with administration of fertility medication to stimulate a woman's ovaries to produce several follicles. Fertilization may occur in the laboratory (IVF) or in the womb.

[2] See, for example, the evolution of success rates in the 2005 CDC Assisted Reproductive Technology (ART) Report at http://www.cdc.gov/ART/ART2005/section5.htm.

[3] In Europe, some countries such as Belgium, Denmark, France, Greece, Israel, Slovenia, and Sweden have complete public coverage for infertility treatment (IFFS Surveillance 07). The U.S. case is examined in this paper.

[4] The prevalence of multiple births is approximately 31% in ART cycles using fresh non-donor eggs or embryos (CDC, 2007) compared to slightly more than 3% in the rest of the U.S. population.

[5] The Total fertility rate for the 25 countries of the European Union is now only 1.5 births (Ziebe and Devroey, 2008) per woman.

in fertility amongst the least fertile women. Typically, these are relatively old women who delayed motherhood and would likely not conceive otherwise (Buckles, 2005; Schmidt, 2005 and 2007). In addition, the greater access to ART has increased the frequency of multiple births in the population (Bundorf *et al.*, 2007). This short-term effect is non-strategic and may be referred to as ex-post moral hazard. The effect of ex-post moral hazard on the average age at first child is a priori ambiguous. The average age at first child increases because infertility is most prominent among older women, who can now extend their reproductive life. In contrast, infertility coverage reduces the benefits involved in waiting for natural conception, thus encouraging women to undergo infertility treatment early, even in situations where a pregnancy could be achieved naturally (Hamilton and McMannus 2005; Bundorf *et al.*, 2007). This effect, would tend to reduce women's age at first birth. The results of previous studies suggest that the former effect prevails, so an increase in women's age at first is expected in the short-term.

In the long-term, however, another result of the policy may occur. In response to easier access to infertility treatments and the possibility to extend the reproductive life, women may be induced to put off motherhood even later. This response by relatively young women, which may be referred to as ex-ante moral hazard, is strategic and would lead to an increase in the average age at first birth several years after the policy was implemented. An increase in the average age at first birth in the mid to long-term would also be consistent with a scenario where initial unmet demand for treatment is gradually satisfied through the opening of fertility clinics throughout the U.S. and/or lower prices. The available data does not enable the empirical assessment of the relative importance of these non-mutually exclusive stories.

The perception that ART increases fertility has led the European Parliament to call on member states to insure the right to universal access to infertility treatment (Ziebe and Devroey, 2008). This movement also is being followed in the U.S. with several attempts at approving the "Family Building Act of 2009," which would extend the coverage for infertility treatments. Although fertility rates may increase in the short-term, they may actually decrease in the long-term if women delay motherhood because of overly optimistic perceptions about their fertility and the effectiveness of infertility treatments (Lampi, 2006; and Benyamini, 2003). The second objective of this paper is to determine whether or not increasing coverage for infertility treatments has an impact on women's life-time total fertility. Even though mandates may negatively affect the total number of biological children per woman through increased delay of motherhood, this effect may be offset through a larger number of children per delivery.

In the U.S., several states, starting with Maryland in 1985, enacted infertility insurance coverage laws forcing health insurance companies to cover infertility treatments to different extents. Several studies have been done on the effects of these infertility mandates on utilization of infertility treatments and other outcomes (Buckles, 2005; Hamilton and McManus, 2005; Schmidt, 2005 and 2007; Bitler, 2005; Bundorf *et al.*, 2007; Bitler and Schmidt, 2011; Mookim *et al.*, 2008). Most of the studies find either direct or indirect evidence of an increase in the usage of infertility treatments after the enactment of infertility insurance mandates, especially for older women. Thus, a well-documented short-term, non-strategic impact of infertility mandates exists for this age group. In addition, Buckles (2005) is the first one to address the impact of infertility mandates on the timing of motherhood and on women's labor market outcomes. She finds that, relative to control states, the birth rates for younger women decreases in mandated states and their labor market participation increases, while that of older women decreases. She interprets these results as supporting the theory whereby mandates allow women to further delay motherhood. However, the question remains whether this reaction by younger women is translated into an age gap between treated and control women that increases over time, or if, prospective mothers learn about the effectiveness of ART and subsequently stop strategically delaying motherhood. In independent and simultaneous work to ours, Ohinata (2009) offers an alternative to Buckles (2005) based on the estimation of a duration model for age at first birth using longitudinal data from the Panel Study of Income Dynamics (PSID). She finds substantial delay of motherhood of approximately $1.5 - 2$ years. Ohinata's identification is, however, based on a relatively small number of women.

The contribution of the present paper to this literature is two-fold: First, it demonstrates that the effect of enacting mandates to cover infertility treatment on the average age at first birth is positive and increases over time. The long-term estimate of the increase in the age of first-time mothers ranges between 3 to 5 months. The synthetic control group method used, developed in Abadie, Diamond, and Hainmueller (2010), relies on more general identifying assumptions than the standard difference-in-differences model and has the additional advantage of assessing how the treatment effect of interest evolves over time. In this part of the analysis, birth certificate data from the National Vital Statistics is combined with the March Annual Social and Economic Supplement of the Current Populations Survey (March CPS) to estimate the effects of infertility mandates on the average age at first birth. Data from the June Marriage and Fertility Supplement of the Current Population Survey (June CPS) also is used to explore the possibility that strategic delay of motherhood in the mid- to long-term is one factor affecting the increase in the average age of first birth. Second, the use of data on the number of biological children (also from the June CPS), shows that mandates do not increase the total fertility rates of women by the end of their reproductive lives. In fact, they tend to reduce the total number of children, although this effect is generally insignificant. This is the first paper to try to estimate the impact on total fertility.

The rest of the paper is structured as follows: Section 2 describes the characteristics of infertility mandates, where and when they were enacted; Section 3 describes the data sources used in this paper; Section 4 presents some motivational statistics and trends as well as the main results about the impact of mandates on the age at first birth; Section 5 presents an analysis of the impact of the mandates on women's fertility over their reproductive lives; Section 6 contains some robustness checks; Section 7 presents conclusions; and Section 8 contains figures and tables. Section 9 is the Appendix.

2 Infertility Treatment Mandates

Table 1 summarizes the main features of infertility insurance mandates and their timing. The classification of mandates in Table 1 is consistent with those presented in Buckles (2005) and Schmidt (2007). Mandates can either require mandatory coverage of infertility treatment for all plans ("mandates to cover") or demand that insurance companies offer at least one plan which covers infertility treatment ("mandates to offer"). In addition, mandates to cover are "strong" when they cover IVF treatment and at least 35% of the women are affected by the mandate, otherwise they are "weak." [6] According to the American Society for Reproductive Medicine, of the six states classified as "mandate-to-cover-strong" only Arkansas does not apply the mandate to all plans (HMOs are exempt). In addition, out of the six strongly treated states, three require women to be married to benefit from the insurance coverage (see Mookin *et al.*, 2008 for more detail on mandates).

Other authors, such as Hamilton and McManus (2005), Bundorf *et al.* (2007), and Mookim *et al.* (2008), classify some states, namely Massachusetts, Illinois, and Rhode-Island (IL-MA-RI), as having "universal," "comprehensive," and "most comprehensive coverage", respectively. In this paper, the effects of infertility mandates for this specific group of states is also analyzed.

The Appendix describes state-specific changes made to the original strong mandates in later periods. Since most of the revisions that occurred within the sample period (*i.e.*, before 2001) undercut benefits, they are expected to decrease the estimated effects of the mandates.

[6] Contrary to Schmidt (2007), Buckles (2005) reports Ohio as a non-IVF coverage mandate.

Table 1: Infertility treatment mandates classifications.

State	cover/offer	mandatory IVF coverage	application	marriage
Arkansas	**cover-strong (1987)**	**yes**	**HMOs excluded**	**yes**
California	offer (1989)	no	All plans	no
Connecticut	offer (1989)	yes	HMOs excluded	no
Hawaii	**cover-strong (1987)**	**yes**	**All plans**	**yes**
lllinois	**cover-strong (1991)**	**yes**	**All plans**	no
Maryland	**cover-strong (1985)**	**yes**	**All plans**	**yes**
Massachusetts	**cover-strong (1987)**	**yes**	**All plans**	no
Montana	cover-weak (1987)	no	HMOs only	no
New York	cover-weak (1990)	no	HMOs excluded	no
Ohio	cover-weak (1991)	no	HMOs only	no
Rhode-Island	**cover-strong (1989)**	**yes**	**All plans**	no
Texas	offer (1987)	yes	All plans	yes
West Virginia	cover-weak (1977)	no	HMOs only	no

Sources: Buckles (2005), Schmidt (2007) and the National Infertility Association (http://www.resolve.org/). Note: Louisiana and New Jersey enacted infertility mandates in 2001, but these states were excluded from our analyses.

3 Data Sources

Data from three main sources is used: 1) birth certificates from the National Vital Statistics System of the National Center for Health Statistics; 2) the March Annual Social and Economic Supplement of the Current Populations Survey (March CPS);[7] and 3) the June Marriage and Fertility Supplement of the Current Population Survey (June CPS).[8]

The analysis of the impact of the mandates on the timing of first births in Section 4 combines data on the timing of first births from the birth certificate data with socioeconomic characteristics available from the March CPS. The birth certificates contain individual records on 50% of the births occurring within the United States during 1968 − 1971; from 1972 to 1984, data is based on a 100% sample of birth certificates from some states and on a 50% sample from the remaining states, and, as of 1985, the data

[7] We downloaded March CPS data and documentation from the IPUMS-USA database (King *et al.*, 2010).

[8] We used processed June CPS files from Unicon Research Corporation (www.unicon.com).

cover every birth from all reporting areas.[9] These data also contain information on the mother, including age, race, and state of residence as well as specific information about the timing, parity (whether it was a first or subsequent birth), and plurality (the number of children per delivery, that is, whether it was a single, twin, triplet, or higher order birth) of each birth. This information allows identification of first births and, therefore, also the determination of the average age of new mothers, which is a variable of interest. When multiple births occur, only one observation per delivery was kept to avoid oversampling multiple-birth mothers, who are more likely to be older and/or to have used ART.[10]

The natality data also contain other potentially relevant socioeconomic variables, such as marital status and maternal education, but the information is not always complete and/or available throughout the sample period.[11] This is why, for the multivariate analyses, the birth certificate information on the age of new mothers is aggregated at the state and year level and combined with a richer set of socioeconomic characteristics obtained from the March CPS, including race, education, marital and labor market status, wages, and health insurance coverage. Note that controlling for employment-sponsored health insurance coverage is important in this context given that uninsured individuals are not directly affected by the mandates and most non-elderly insured individuals in the U.S. obtain insurance through their workplace.[12]

Our analysis could be conducted only until 2005 because after that year the natality data lacks state identifiers. However, our study was restricted to the period before 2001 (*i.e.*, from 1972 to 2001)

[9] Births occurring to U.S. citizens outside the United States are not included. The number of states from which 100% of the records are used increases from 6 in 1972 to all states and the District of Columbia in 1985. We adjusted the total numbers accordingly in the analysis.

[10] We uniquely identify multiple-birth mothers by using, whenever available, various variables such as year, month and day of birth, gestation time, and state, county and place or facility of birth, presence of attendant at birth, plurality, maternal age, race, years of schooling, marital status, place of birth, and state, county, city, and standard metropolitan statistical area (SMSA) of residence and paternal age and race.

[11] Importantly, information on maternal education is missing for the following states and years: California (1972 – 1988), Alabama (1972– 1975), Arkansas (1972 – 1977), Connecticut (1972), District of Columbia (1972), Georgia (1972), Idaho (1972 – 1977), Maryland (1972 – 1973), New Mexico (1972 – 1979), Pennsylvania (1972 – 1975), Texas (1972 – 1988) and Washington (1972 – 1991). Marital status is not reported in any state until 1978.

[12] An important feature of state-mandated benefits is that self-insured employers are exempt from state insurance regulations under the 1974 Federal Employee Retirement Income Security Act (ERISA). Hence, employers who self-insure are exempt from the requirements of the state infertility insurance mandates previously described. Since self-insured companies are typically large, the impact of the mandates is likely to be concentrated on small firms. Lacking information on the self-insured status of employers, researchers have used firm size as a proxy for ERISA exemptions (*e.g.* Schmidt, 2007 and the references therein, Simon, 2004, Bhattacharya and Vogt, 2000). Self-reported firm-size from the March CPS could be used as a proxy for ERISA exemption status but unfortunately this variable was not recorded before1988 and therefore could not be included as a predictor in our estimations.

so that Louisiana and New Jersey could be included as controls (these two states passed infertility insurance laws in 2001.) Including Louisiana and New Jersey in the treated group would not have provided us enough post-intervention years to analyze the long-term impact of these latest mandates. In addition, since states are not uniquely identified in the CPS until 1977, the analyses could be enriched by incorporating March CPS variables only from 1977 onwards. To further enrich the set of control variables, state-year legal abortion rates by 1000 women aged $15-44$ and state of residence obtained from The Guttmacher Institute were included.

In our analysis of the total number of biological children born to women of childbearing age in Section 5, data from the June CPS is used. Unlike the March CPS, which is available on a yearly basis and only provides information on the presence of children in the household without discriminating between biological and non-biological children, the June questionnaire is not administered every year but contains information on the number of biological children ever born. In particular, the June CPS provides this information for the following years during our sample period: $1979-1985$, $1990-1992$, $1994-1995$, 1998, and 2000.[13] Additionally, the June CPS contains information on other potential determinants of fertility, such as age, marital status, and labor market status which were incorporated as controls in the regressions.

4 The Effect of Infertility Mandates on Average Age of First Birth

4.1 Descriptive Evidence

Figures 2a, b, and c plot the evolution of the age of new mothers in control states *versus* all treated states, all strongly treated states, and Massachusetts, Illinois and Rhode Island, respectively. The two vertical lines in each figure indicate the years in which the first and last of the corresponding mandates were passed; $(1977, 1991)$ for all the treated states, $(1985, 1991)$ for the strongly treated states and $(1987, 1991)$ for Massachusetts, Illinois and Rhode Island. While the average age of first-time mothers was higher in treated than in control states even before any mandate was enacted, Figures 2b and c show that for states with "strong mandates to cover" and for Massachusetts, Illinois, and Rhode Island, the treated-control gap became larger after the passage of the mandates. As expected, this trend is not

[13] In 1977, 1986, 1987, and 1988 the "number of babies" question also was asked but only to women who had ever been married. The question is most often posed to women in their childbearing years, which in the June CPS was usually to women ages $18-44$. Including women aged $45-49$ would limit the analysis to the years 1979, 1983, 1985, and 1995, leaving us with very few post-intervention periods.

so evident when all of the treated states were considered together (Figure 2a), given that both the "weak mandates to cover" and the "mandates to offer" are much more limited than the "strong mandates to cover."

[Figure 2 here]

More specifically, in 2001, 10 years after the last strong mandate passed in Illinois, the age gap between strongly treated and control states is slightly larger than 1 year, which constitutes an increase of 0.42 years (or 5 months) with respect to the size of the gap in 1985, when the first strong mandate passed in Maryland. This double difference represents 25% of the overall increase in the age of new mothers that occurred in strongly treated states between 1985 and 2001, and is consistent with previous studies showing that the enactment of infertility treatment mandates led to an increase in birth rates for women older than 35 (Schmidt, 2007; Buckles, 2005). This pattern also is present if the analysis is restricted to White new mothers only (Figure 3), for whom the corresponding double difference figure in percentage terms also amounts to approximately 25%. The corresponding values for Massachusetts, Illinois, and Rhode Island are 31% and 32% for all new mothers and for White new mothers, respectively.

[Figure 3 here]

Particularly relevant to this analysis is the widening of the gap in average age at first birth between mandated and non-mandated states several years after the last mandate passed, also visible in Figures 2 and 3. This increase suggests that the long-term cumulative impact of the mandates on the timing of first birth is likely to have gone beyond a short-term impact on older women with infertility problems whose access to ART was facilitated by the mandates. Conceivably, the passage of infertility mandates may have induced a behavioral response among younger cohorts of women whose childbearing decisions were further delayed because of the lower cost of ART. This behavioral response or strategic effect may be reinforced if women's childbearing age is affected by the age at which their peers' have babies.

Other reasons for the increasing effect of the mandates may be related to growing access to ART for those who are not directly affected by the mandates. As Hamilton and McManus (2005) show, the enactment of some mandates brought about an increase in the average size of fertility clinics, which most likely allowed lower price-cost margins potentially benefitting all patients including those without coverage for these treatments. Empirically assessing the relative importance of these two non-mutually exclusive stories is, however, beyond the scope of this paper. In Section 4.3, however, some evidence is presented that suggests the behavioral response previously described may have played a role.

To assess whether the age difference between new mothers in treated and control states grew over time within a regression framework that takes into account state- and year-specific factors, the following simple model was estimated:

$$\begin{aligned} Mage_{ist} &= \alpha + \beta Yearscov1\_5_{st} + \gamma Yearscov6\_10_{st} + \delta Yearscovmore10_{st} \\ &+ \sum_s \theta_s S_s + \sum_t \lambda_t Y_t + \varepsilon_{ist} \end{aligned} \tag{1}$$

where the dependent variable, $Mage_{ist}$, is the age at first birth of woman i in state s and year t. Time-invariant state-specific factors that affect the timing of motherhood are captured by state fixed-effects, S_s, while Y_t denotes year fixed-effects that capture trends in the timing of first births common to all women across the nation. The independent variables of main interest are a set of indicators for whether new mother i gave birth in a state s when an infertility insurance mandate had been in place for 1 to 5 ($Yearscov1\_5_{st}$), 6 to 10 ($Yearscov6\_10_{st}$) or more than 10 years ($Yearscovmore10_{st}$). Notice that Eq. (1) does not incorporate any socioeconomic characteristics because the birth certificate data lack most of them. Finally, ε_{ist} is a mother-specific error term, capturing all purely idiosyncratic factors that influence the timing of first births.

Table 2 displays OLS estimates of β, γ and δ from equation (1). The first two columns display the results of estimating equation (1) for all new mothers and White new mothers for all states. The two middle columns estimate (1) excluding all births occurring in states with "mandates to offer" and "weak mandates to cover" restricting treatment to strong coverage mandates, while the last two columns further restrict our set of treated states to those with the "most comprehensive coverage", that is, Massachusetts, Illinois and Rhode Island.

[Table 2 here]

Not surprisingly, given the nature of the "mandates to offer" and the "weak mandates to cover" described in Section 2, coefficient estimates (reported in columns 1 and 2) obtained when considering all treated states are small in magnitude, lack significance at standard levels of testing, and sometimes are even negative. The picture changes completely when focusing on the impact of strong coverage mandates and excluding all the other treated states from the sample in columns 3 and 4. Then, the variables of interest, $Yearscov1\_5_{st}$, $Yearscov6\_10_{st}$, and $Yearscovmore10_{st}$ are jointly statistically significant, and their estimated coefficients clearly indicate that the impact of strong mandates increases over time, with the differences across the corresponding coefficients statistically significant. As expected, this pattern is reinforced when concentrating on Massachusetts, Illinois and Rhode Island.

Although the evidence presented in Table 2 is very suggestive, to interpret it causally the control group should be valid, *i.e.*, it should be able to reproduce what would have happened in the treated group had the mandates never passed. This issue is addressed in the next subsection.

4.2 Synthetic Control Group Estimates

The traditional way to estimate the effect of the infertility mandates on women's age at first birth would be to rely on a difference-in-differences model (DID) such as Eq. (1), usually augmented with a set of control variables. DID estimators are often used to evaluate the impact of policies or interventions that affect aggregate units (*e.g.*, states in this paper). The basic idea behind the DID estimator is to compare the evolution of the outcome of interest for units affected by the intervention (treated units) with the evolution of the same outcome for unaffected units (control units). Identification requires the average outcome for the treated unit to experience the same variation as the average outcome for the controls in the absence of the intervention. This restriction may be implausible when the distribution of characteristics that are likely to influence the evolution of the outcome variable differs between treated and controls units. Researchers usually address the latter issue by incorporating a rich set of covariates into a regression framework.

To construct a control group that maximizes the similarities between women in treated and control states, the synthetic control method recently developed by Abadie, Diamond and Hainmueller (2010, henceforth ADH),[14] is used, which presents several advantages over the conventional DID estimator. The synthetic control group approach limits the discretion of researchers in the choice of the control units by offering a procedure for the construction of an "ideal" control group, which they denote as "synthetic" control group. The synthetic control group uses a weighted average of the potential control units, which provides a better counterpart for the treated units than any single actual control unit or a set of actual control units. The weights assigned to each control unit are chosen to minimize the differences in pre-treatment trends and exogenous regressors, denoted by "predictors" in ADH, between the treated unit and the synthetic control group. This estimation procedure is very transparent since

[14] See Abadie and Gardeazabal (2003) for an earlier application of the synthetic control group approach.

the relative contribution of each control unit to the synthetic group, which may be zero, is made explicit.

It is worth noting that, while the synthetic control group approach is obviously related to the standard DID estimator, which it extends, it also has features in common with matching estimators since both approaches attempt to minimize observable differences between the treatment and control units. Indeed, some of the latest developments in the literature attempt to minimize the chances of selection into treatment based on unobservables.[15] The synthetic control approach is a step in this direction since it relies on more general identifying assumptions than the standard DID model, allowing the effects of unobserved variables on the outcome to vary with time.

To apply the synthetic control group, the birth certificate data on the age of new mothers must be aggregated at the state and year level. This aggregation is advantageous because it allows us to control for socioeconomic characteristics by merging the birth certificate data with socioeconomic variables available in the March CPS (also aggregated at the state and year level). All births from strongly treated states also are aggregated and use 1985, the year the first strong mandate was enacted, as the initial treatment year. Additionally, information for the states with the most comprehensive mandates, Massachusetts, Illinois, and Rhode-Island, also was aggregated and the analyses is replicated for this subset of treated states, where the first mandate was enacted in 1987 in Massachusetts.

The synthetic control group is constructed as the convex combination of control states (see Table 3 for the estimated weights) that are most similar to the states with strong coverage and comprehensive coverage in terms of various socioeconomic predictors as well as lagged values of average age of first motherhood before treatment (*i.e.*, before 1985). More precisely, the predictors chosen include: 1) variables that control for the demographic and family structure of the female population, such as the percentage of new mothers older than 35, and the percentage of married women in the state; 2) variables that control a state's race composition, such as percentage of white and black females; 3) variables that control for the education level of the female population, such as the percentage of highly educated women; 4) variables related to the female labor market, such as the participation rate and employment rate, the average logarithm of the hourly wage, and the percentage of women covered by Employment Sponsored Insurance (ESI); 5) the per 1000 women abortion rate by state of residency; and 6) several

[15] These concerns were raised in several studies (*e.g.* Heckman *et al.*, 1997, Heckman *et al.*, 1998, Michalopoulos, 2004, Smith and Todd, 2005) where it was argued that matching on observables alone would not guarantee an adequate counterfactual because unobservables may affect the selection into treatment leading to bias in the estimation of treatment effects. Heckman *et al.* (1997), Heckman et al. (1998), and Smith and Todd (2003) present evidence that highlights the advantages of using a DID matching strategy, which allows for time-invariant differences between the treatment and control groups. Michalopoulos (2004) allows for selection into treatment based on individual-specific unobserved linear trends.

lags of average age at first birth.[16] All these predictors are averaged over different periods to maximize the fit of the estimation. Although the predictors are roughly the same for the four estimations (strong, strong White only, IL-MA-RI, IL-MA-RI Whites only), the composition of the synthetic control group is not the same as can be seen in Table 3. The most important state in the composition of the four synthetic control groups is New Jersey, which represents between 26 to 41% of the estimated synthetic control group.

[Table 3 here]

Table 4 displays the pre-treatment (*i.e.* before 1985) sample averages of all predictors for the states with strong coverage (column 1), as well as for the synthetic control group (column 2), and for the full group of control states (column 3). Table 5 replicates Table 4 but focuses on the sample of White women. Finally, Tables 6 and 7 are equivalent to 4 and 5 except for Massachusetts, Illinois, and Rhode-Island as the treated group. As can be seen in all four tables (4, 5, 6, and 7), prior to the passage of the first strong mandate to cover, new mothers already were clearly younger in the control states than in states where strong mandates to cover eventually passed. They also earned lower wages on average, were less educated, more likely to be married, less likely to abort, less likely to participate in the labor market, and less likely to be employed and to have employer-provided health insurance coverage. The predictors' pre-treatment values for the strongly treated states and the subset of Massachusetts, Illinois, and Rhode-Island, as shown in Tables 4, 5, 6, and 7, resemble the pre-treatment values of the synthetic control group (column 2) much more than the pre-treatment values for the full set of control states (column 3). Hence, the synthetic control group should be a better counterfactual for the treated groups.

[Table 4 here]

[Table 5 here]

[16] Other variables were considered as predictors but were discarded because they worsen the fit of the model [*i.e.* they increased the root mean squared prediction error (*rmspe*) of the estimation, which is a measure of the difference between the treated and the synthetic control group during the pre-treatment period]. These were, for example, the average number of children in the household, the split of the female population's age structure into 5-year age brackets, the percentage of females with private health insurance, percentage of first-delivery at different 5-year age brackets, average company size for female workers, and the year of divorce reforms according to Friedberg (1998) and Gruber (2004).

Our synthetic control estimate of the impact of the infertility coverage mandates on the timing of the first child is the difference between the average age of new mothers in states with "strong mandates to cover" (and the subset of MA-IL-RI) and the synthetic control group. The first panel of Table 8 shows estimates of the effect for the group of states with strong mandates to cover while the second panel of the table shows the same estimates for the subset of these states with the most comprehensive mandates (IL-MA-RI). The second column of Table 8 reports the synthetic control group estimate of the effect of strong mandates in 2001, that is, 16 and 10 years after the first and the last strong mandates were passed, respectively. We refer to this estimate as the long-term effect of the mandates. For the group of states with strong mandates, the long-term effect amounts to 0.266 and 0.317 years, roughly 3.2 months for all women and 3.8 months for White women, respectively. For IL-MA-RI, as predicted, the effects are larger although the number of years since the first mandate is lower. These values amount to an increase of roughly 4.1 to 5.4 months in the average age at first child for all and for White new mothers, respectively.

This long-term effect of the strong mandates is sizable—between 15.7% and 18.8% of the total increase from 1985 to 2001 for the group with strong coverage and between 24.8% and 34.3% for IL-MA-RI. The synthetic control estimate also is slightly less than the one obtained from the raw DID aggregate estimate (Table 8) which is approximately 0.42 years (5 months), even less than the value for the individual level DID estimates presented in Table 2 in Section 4.1.

[Table 8 here]

The p-values shown in column 3 of Table 8 are computed using the inferential method, proposed by ADH, to construct confidential intervals. The method assigns treatment to each of the control states and estimates what ADH denote as "the placebo treatment effect" for each of the 38 control states. The idea is that the placebo treatment effects should be close to zero. The p-value indicates the real treatment effect in the distribution of all estimated effects ranked according to size. Therefore, a p-value of 0.158 for the estimated effect for the strong mandates for all samples indicates that it is within the 15.8% of the largest effects (including the real effect). None of the estimated effects is statistically significantly positive for standard significance values, according to the p-values shown in column 3. However, obtaining high p-values is common when using the placebo tests method because it may not be possible to find a good synthetic group for some of the control states (*i.e.,* if a state is extreme, then the fit of the pre-treatment predictors may be poor). When this is the case, ADH recommends discarding such placebo treatments for the purpose of computing the p-values. This is precisely what is done in column 4 of Table 8. The "p-value5" is constructed by discarding all those placebo treatments with a fit worse than five times the root mean squared prediction error (*rmspe*) of the real treatment. The *rmspe* is a measure of the difference in age at first birth between the treated and the synthetic control group during the pre-treatment period. Hence, the lower the *rmspe*, the better the model fits the data. The *rmspe* values, displayed in column 5 of Table 8 are remarkably low compared to those obtained for the placebos. According to p-value5, the effects would be statistically significant at 9.4 and 9.1% levels for the entire sample of women for strong mandates and for IL-MA-RI. Although the effects are larger for White women, they are not statistically significant at the 10% level although the p-values are low.

Another way to assess the significance of the treatment effect is by looking at the size of the post-pre ratio of the *rmspe* for the treated states relative to the placebos (ADH). If there is no treatment effect, the ratio of the post-pre *rmspe* should be approximately the same for the treatment units as for the placebos. The values of the post-pre treatment *rmspe* ratio for the strong treated states were 10.49 and 11.08 for all women and White women, respectively, and 7.36 and 8.72 for all women and White women, respectively, for the subset of IL-MA-RI. The p-values for these ratios, shown in column 6 of Table 8, are all very low, indicating that the ratios are all statistically significantly different from zero. One advantage of this test is that, unlike the previous one, it takes into account all control states, and therefore eliminates the need for arbitrary choices regarding which placebo estimates should be discarded.

Figure 4 shows the annual average age at first birth in strongly treated states and in IL-MA-RI compared to the synthetic control group counterpart for the sample period ($1972 - 2001$) for all women and White women. The synthetic control group does a good job in tracking the pre-treatment evolution of new mothers' age in states with strong coverage and in IL-MA-RI, which indicates it is a good approximation to the counterfactual trend in maternal age at first birth that states with strong coverage would have experienced had the mandates not been enacted. This is not surprising, given the closeness in terms of predictor values between the states with strong coverage and their synthetic version shown in Tables 4, 5, 6, and 7, and is also consistent with the very low values obtained from the *rmpse*.

[Figure 4 here]

More important than the size of the estimated long-term effect of the strong mandates is its evolution over time, which is shown in Figure 5. This shows an increase in the effect of the strong mandates over time. Regressions of the estimated annual effects of the 17 post-treatment periods for the strong mandate states (and 15 post-treatment periods for MA-IL-RI) on indicators of time since the mandates (*i.e.,* less than 5 years since the mandate, between 6 and 10 years, or more than 10 years), shown in Table 9, confirm that the impact of the mandates grew significantly over time. The long-term cumulative impact of the mandates on the timing of first births, therefore, goes beyond its short-term impact on older women with infertility problems. We believe the mechanism operating here is simple. Suppose no supply constraints existed for infertility treatments when the mandates were enacted. Then, if mandates had only a non-strategic effect on older women (*i.e.,* ex-post moral hazard), the estimated effect should be positive but nearly constant over time. The long-term effect may be larger than the short-term because, for example, women who were young when the mandates were enacted could strategically delay motherhood (*i.e.,* exert ex-ante moral hazard). An alternative explanation for the increasing effect may be that supply constraints for fertility treatments existed when the mandates were enacted, but the response from the supply side (*e.g.,* technological improvement and/or price reductions) was able to absorb a larger number of users of infertility treatments. Our data cannot identify the exact

contribution of each of these potential explanations to the increasing effect of the mandates, but this is discussed further in the next Section.

[Figure 5 here]

[Table 9 here]

4.3 Discussion and Interpretation of Results

In this Section, we first provide some evidence indicating that part of the explanation for the growing gap may be due to strategic delay of motherhood. Second, we discuss the possibility that welfare reform legislation, may be affecting the results.

The sizable and increasing effect of the strong mandates on average age of first birth, although suggestive of an increase in strategic delay, does not prove it. A plausible alternative explanation for the growing effect over time is related to the potential supply response to the mandates as mentioned at the end of the last Section. In this Section, support is provided for the theory of strategic delay by estimating the effect of time since enactment of the mandate on the probability of having at least one biological child by age 30 and 35 using data from the June CPS.[17] In these regressions, women are by definition older than 30 or 35 years. We first estimate the impact of the number of years of mandated coverage at the time of the interview. The number of years with mandated coverage at the time of the interview may not, however, reflect the number of relevant years "under treatment." For example, a 44-year-old woman who had her first child when she was 20, and has been under the mandated coverage for 10 years, did not have, by definition, her probability of having a biological child before the age of 35 years affected by the mandate. To ameliorate this measurement error and better reflect the intention to treat, we use the variables "number of years of mandated coverage at age 30" and "number of years of mandated coverage at age 35". Note that under these new definitions, the woman in the example would have zero years of mandated coverage when she was 30 years old although she still shows one year of mandated coverage by the age of 35. In addition, age interval dummies are added to all of the regressions.

Panel A of Table 10 shows the estimated marginal effects of the number of years of mandated coverage at age 30 and 35 on the probability of having at least one child at that age. Panel B shows the effect of number of mandated years at the time of the interview on the probability of having at least one child by 30 and 35 years. The marginal effects are obtained from probit estimations done for all women and for White women only. A large set of controls are included, such as state fixed effects, year fixed effects, and age dummies of 5-year intervals (see note in Table 10 for a complete description). In a given state and year, the number of years of mandated coverage by 30, varies by women according to age. For example, a woman from Maryland who turned 30 before 1985 has zero years of mandated coverage by 30 whereas a Maryland woman who turned 30 in 1990 has 5 years of mandated coverage at age 30, and 10 years of mandated coverage by 30 if she turned 30 in 1995. Therefore, the coefficient on $1-5$ years of mandated coverage at age 30 is being identified by relatively older women, while the $6-10$ years

[17] To construct the variable "at least one child by age x," where we use $x = 30$ and 35 years of age, we need to know the age at which each woman has her first child. The latter is constructed from the June CPS's variable *birth1y* which reports the year of birth of the first child. Unfortunately, this variable is missing for all states in years 1984, 1994, 1998, and 2000. In addition the June CPS is not available for the years 1977, 1978, $1986-1989$, 1991, 1993, 1996, 1997, 1999, and 2001. This implies that we cannot construct the variable "at least one child by age x" beyond 1995.

For the years 1990, 1992, and 1995 *birth1y* is missing many values but they are spread across all states. The number of missing values for *birth1y* is approximately 53% and 55% for control and strongly treated states, respectively. Note that the June CPS data is more suitable than the March CPS for this analysis because the latter does not have information on the number of biological children but instead provides the number of children (which includes adopted children, for example) in the household (and therefore children who do not live in the household are not included). Presumably, for relatively young women (whose children have not left the household and whose probability of adoption is smaller) the information on the March CPS would work as well as the June CPS but this would not be true for older women.

of mandated coverage at age 30 is being identified by the younger cohorts. Note that because states enacted their mandates in different years, the number of years of mandated coverage by a certain age is not colinear with age; *e.g.,* a woman experiencing 5 years of mandated coverage by age 30 in Illinois is 6 years younger than a woman from Maryland with the same years of coverage by age 30.

[Table 10 here]

Results from the first four columns of Table 10 show that having a strong mandate for longer than 6 years is associated with a significant lower probability of having a child by the age of 30. The marginal effects means a reduction of about 1.9 percentage points to just above 3.5 percentage points in the probability of having a child by the age of 30. These effects are smaller in magnitude for White women and, except for one, are not statistically significant by the age of 35 implying that by then most women decide not to delay further. Although mostly still negative, the effects of $1-5$ years of mandated coverage either by the age of 30 and 35 or at the time of interview are in general statistically insignificant. These results suggest a delay in the timing of motherhood due to the mandates consistent with Buckles (2005) and Ohinata (2009). Moreover, they seem to suggest that younger women, who have had more time to react to the mandates, are using the mandates to strategically delay motherhood.

[Table 11 here]

The exercise was repeated for the comprehensive states Illinois, Massachusetts, and Rhode Island (Table 11). This analysis shows larger (in absolute value) and more statistically significant marginal effects for the $6-10$ years of mandated coverage by age 30 in Panel A. Interestingly, a positive marginal effect is found for $1-5$ years of mandated coverage. These positive marginal effects are consistent with a higher usage of infertility treatments when mandates were enacted (Bundorf et al., 2007 describe the potential moral hazard among relatively fertile couples), which appears offset by an increase in delay by younger women. By age 35, no statistically significance is obtained, indicating that women at that age stopped delaying motherhood.

The regressions presented in Tables 10 and 11 are similar to those presented by Buckles (2005), with some key differences. First, Buckles uses data from the March CPS and the variable she uses for the older women sample is "presence of small children in the household." These data have at least two problems: first, the relatively older women may not have small children in the household because they do not have children or because their children are already grown. Second, the variable does not restrict children to biological children, yet women who delayed motherhood may be more likely to adopt children. Buckles also estimates a similar regression for younger women where the problems just mentioned are less likely to occur. She finds lower prevalence of young children in the household among young women in treated states, which is consistent with the findings reported here (for a slightly older group).

Finally, other potential causes may exist for the growing gap between treated and synthetic states. For example, during the post-treatment years either treated or control states may have enacted laws—*e.g.,* welfare reform—that affected the mean age at first birth. Welfare reform is likely to have discouraged maternity at younger ages through its demanding work requirements and stringent eligibility standards for acceptance into the assistance programs and, therefore, could have increased the mean age at first birth.[18] The welfare reform was enacted in 1996 [Personal Responsibility and Work Opportunity Act (PRWORA)] and became effective in July 1997, 4 years before the end of our sample. Before that, however, some states had already introduced work requirements for welfare eligibility in the 1980s and early 1990s (Meade, 2004). For the interpretation of our analysis results it is important to know the group (treatment, control, or not in the sample) to which these early adopters of welfare reform belong. If some of the treated states were early adopters of welfare reforms, then the present results would likely overestimate the effects of strong infertility mandates on the mean age at first birth. In contrast, if early adopters constitute part of our synthetic group, then the estimated gap in mean age at first birth between treated and synthetic states would be underestimated. Reports show that early adopters [California, Colorado, Iowa, Michigan, Oregon, Wisconsin, and Utah (see Meade, 2004)],— with the exception of California (which is neither a treated nor a control state)—, are control states and, hence, we should expect, if anything, a downward bias in our estimates of the effects of the strong infertility mandates on the average age at first birth.

5 Do Women With More Access to ART Catch Up in Terms of Total Fertility?

Infertility mandates may result in ex-ante moral hazard and cause women to delay motherhood. However, even if ex-ante moral hazard does occur, would that necessarily result in a lower number of children per woman? At least two factors operating here may have opposite effects. Mandates may negatively affect the total number of biological children per woman if the mandates cause a further delay in moth-

[18] To our knowledge, no study has found that welfare reform increased the age of first motherhood. Instead, Hao and Cherlin (2004) compare two cohorts of young women and conclude that welfare reform has not decreased teenage fertility.

erhood. In contrast, any negative effect on the number of deliveries may be compensated by a higher average number of children per delivery since infertility treatments increase the probability of multiple births. In this Section, we estimate the effect of strong mandates on the total number of biological children per woman. The results from this Section contribute to the debate on policies to increase fertility rates in Europe.

Figure 8 presents two cohorts (born between $1949-1952$ and between $1954-1957$) and plots the average number of biological children over a woman's reproductive life for control, strongly treated, and comprehensive states (*i.e.*, Illinois, Massachusetts, and Rhode-Island) using data from the June CPS.[19] For both cohorts, women in strongly treated and comprehensively treated states have on average a lower number of biological children and do not catch up with women in the control groups by the age of 44.

To control for other covariates that may affect the trends shown in Figure 8, we also estimate a zero inflated Poisson regression of the number of biological children against a number of covariates. Tables 13 and 14 show the marginal effects of time since the mandates for a sample of 44-year-old women (*i.e.*, women at the end of their reproductive lives), and all women, respectively. Each of these tables also shows marginal effects for a sample of White women only and for Illinois, Massachusetts and Rhode-Island as the treated group.

The tables show no effect of the mandates in total fertility for the sample of 44-year-olds, not even when the treatment group is restricted to the comprehensive states. When all women are included in the regression, a negative and statistically significant effect is found for longer than 6 years of mandate coverage, but this effect is weaker for the comprehensive states, and disappears once the sample is restricted to White women. Several robustness checks are performed. First, all models were re-estimated using only two time dummies, $1-5$ years of mandated coverage, and more than 5 years of mandated coverage. None of the coefficients is statistically significant for the 44-year-old sample, and similar results are obtained for all women. Second, the model was re-estimated as a linear regression and the results are qualitatively similar. Finally, the model was re-estimated for the 44-year-old sample while restricting the sample to mothers using a Poisson regression, but nothing is found to be statistically significant. It is clear from the exercise that infertility mandates have not increased women's total fertility.

[19] In the June CPS, the number of biological children was obtained systematically only from women who were 44 years old or younger.

6 Robustness Checks

Next, an analysis is conducted to determine whether the enactment of mandates encouraged women who were more likely to benefit from insurance coverage of infertility treatment to move to the enacting states. Married women who are childless after a certain age are more likely to benefit from infertility mandates and, therefore, may move from control states to treatment states. If this population did move, our treatment effects would be biased upwards. Figures 6 and 7 compare the change over time in treatment and control states of the percentage of relatively older women (between $30-49$ and $35-49$ years old) and the percentage of relatively old women who were married and childless. In general, these figures do not show an increase in the percentage of these groups of women in the strong treatment states relative to the control states with the exception of an increase in the percentage of women who were married and childless after 30 years old between 1985 and the early 1990s. This increase soon vanishes and is followed by a sharp decrease in the following years. Since it can be difficult to draw conclusions from figures, the unconditional DID by state is also computed and shown in Table 12 . The unconditional DID estimate is always negative for the strongly treated states. From the group of comprehensive mandated states (MA-IL-RI), the unconditional DID on the percentage of women $30-49$ and $35-49$ is positive although very low, representing 0.7% and 2.5%, respectively, of the values in 1985. However, if we look at the unconditional DID for the percentage of women in those age groups who are married and childless, they are also negative for the group of comprehensive mandate states.

7 Conclusions

This paper poses two questions about the impact of infertility treatment insurance coverage on fertility. First, does the coverage of infertility treatment have an effect on the average age of first-time mothers, and, if so, does it increase over time? Second, does the coverage of infertility treatment increase total fertility by the end of a woman's reproductive life due to the higher prevalence of multiple births, for example? Variation in the enactment of infertility insurance mandates over time and across U.S. states is exploited to answer the two questions. Infertility mandates vary across states in several ways, but essentially can either require mandatory coverage of infertility treatment for all plans ("mandates to cover") or demand that insurers offer at least one plan which covers infertility treatment ("mandates to offer"). In addition, mandates to cover are "strong" when they cover IVF treatment and at least 35% of the women are affected by them, otherwise they are "weak." Infertility mandates have been enacted in some U.S. states mainly during the late 1980s and early 1990s. After confirming the expectation that "mandates to offer" and “weak” mandates were ineffective, we focus our attention on the effect of the strong mandates to cover, which were offered in 6 states (the treatment group): Arkansas (1987), Hawaii (1987), Illinois (1991), Maryland (1985), Massachusetts (1987), Rhode-Island (1989). We combine Birth certificate data from the National Vital Statistics and the March Annual Social and Economic Supplement of the Current Populations Survey (March CPS) to estimate the effect of the infertility mandates over time on the average age at first birth. To estimate the effects of the infertility mandates on women's total fertility we use the Marriage and Fertility Supplement of the Current Population Survey (June CPS).

Results show that the effect of enacting strong mandates to cover infertility treatment is positive on the average age at first birth and increases over time. The long-term estimates of the increase in age of first-time mothers range from 3 to 5 months. The estimation method is the synthetic control group method developed by Abadie, Diamond, and Hainmueller (2010), which relies on more general identifying assumptions than the standard DID model. Results also show that strong mandates do not increase the total fertility rates of women by the end of their reproductive lives, in contrast, they tend to reduce the total number of children, although this effect is generally insignificant. Coverage of infertility treatment has been considered as a potential policy to increase European fertility rates. Our results suggest that such a policy would not contribute to a long-term increase in fertility.

8 Figures and Tables

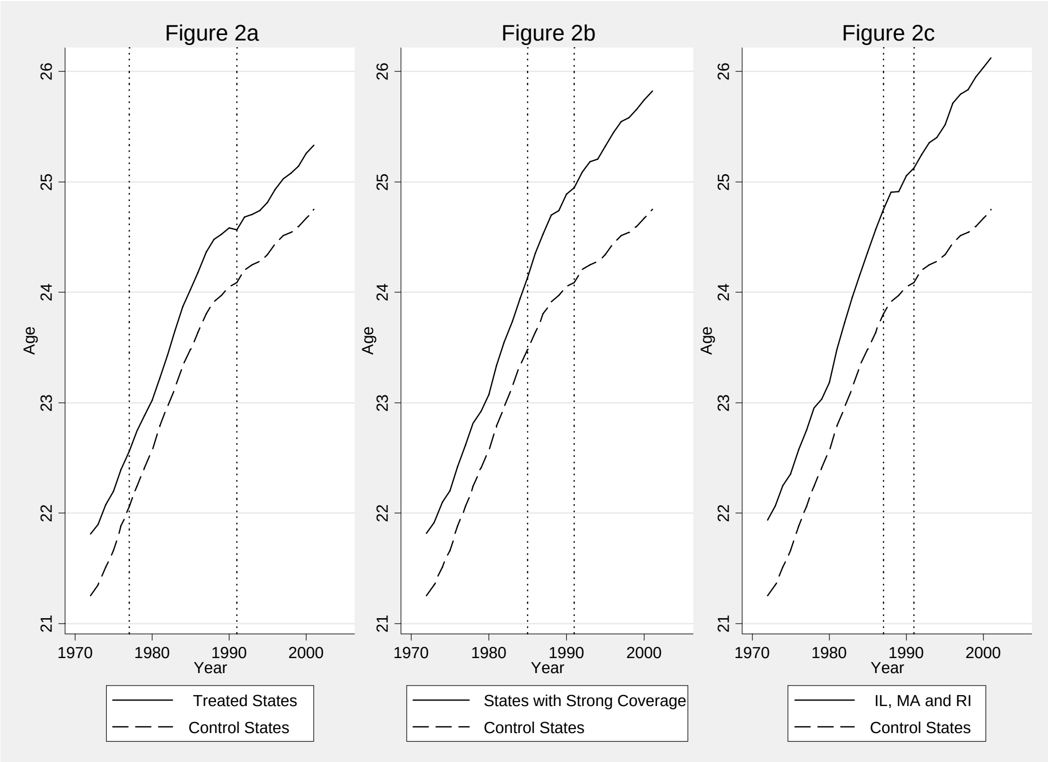

Figure 2: Maternal age at first birth. All women

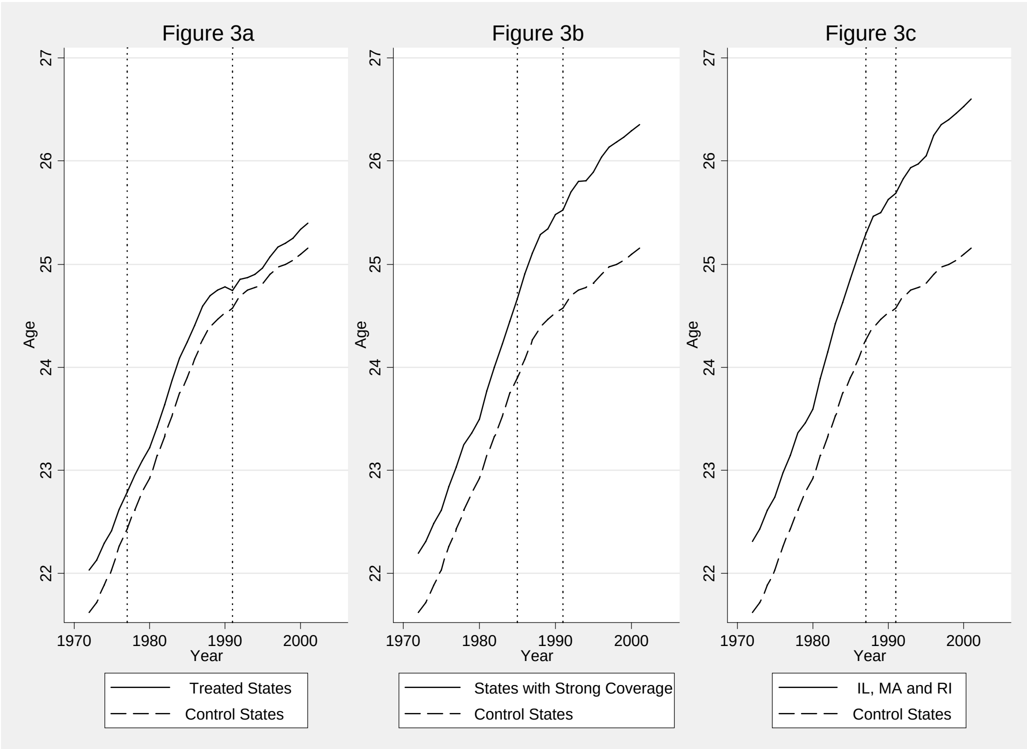

Figure 3: Maternal age at first birth. White women

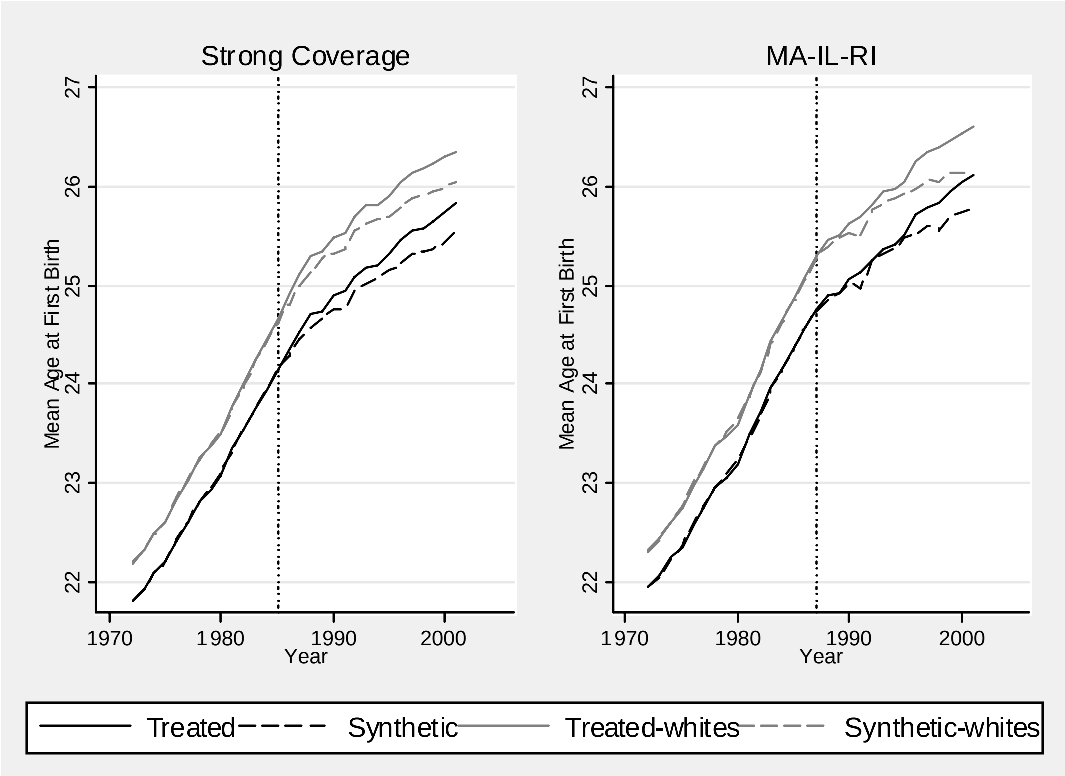

Figure 4: Age at first birth: evolution in treated states and their synthetic control groups

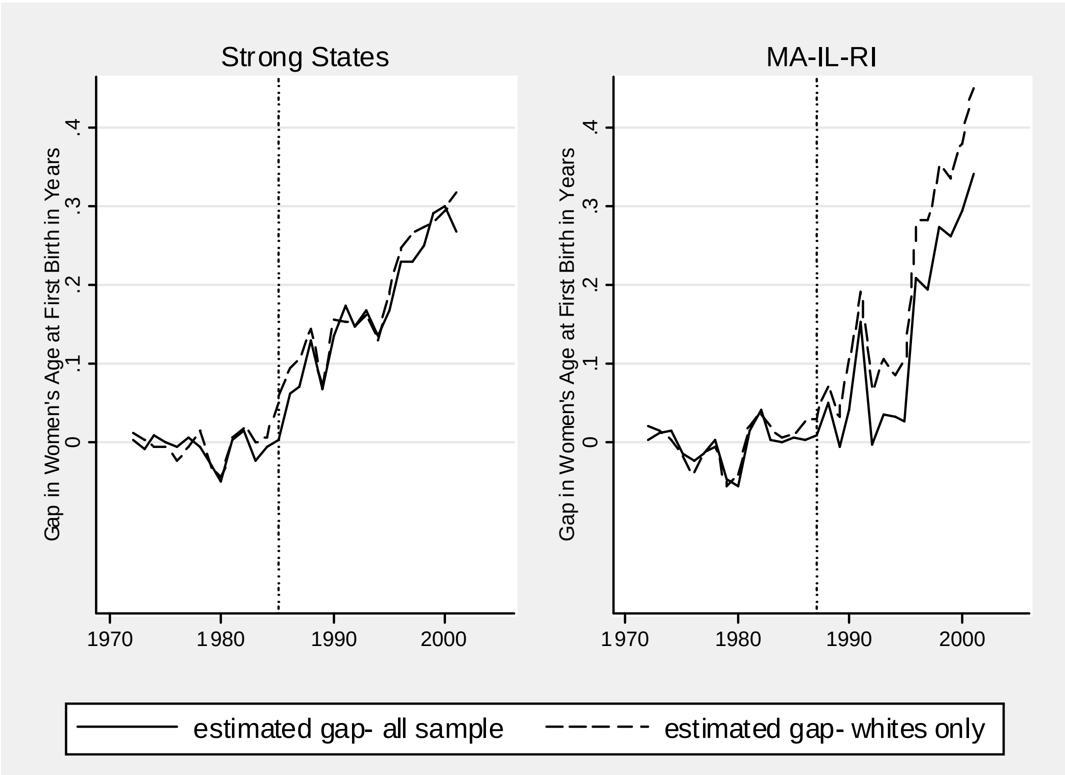

Figure 5: Gap between treated and synthetic groups for age at first birth

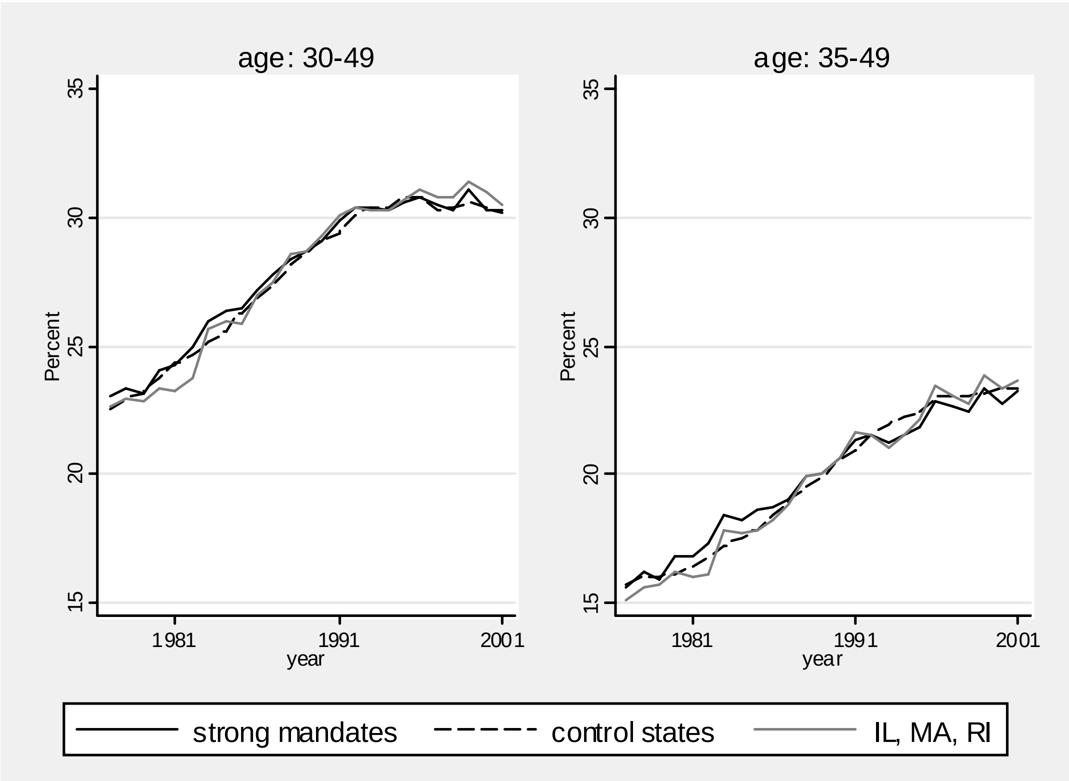

Figure 6: Percentage of women by age group

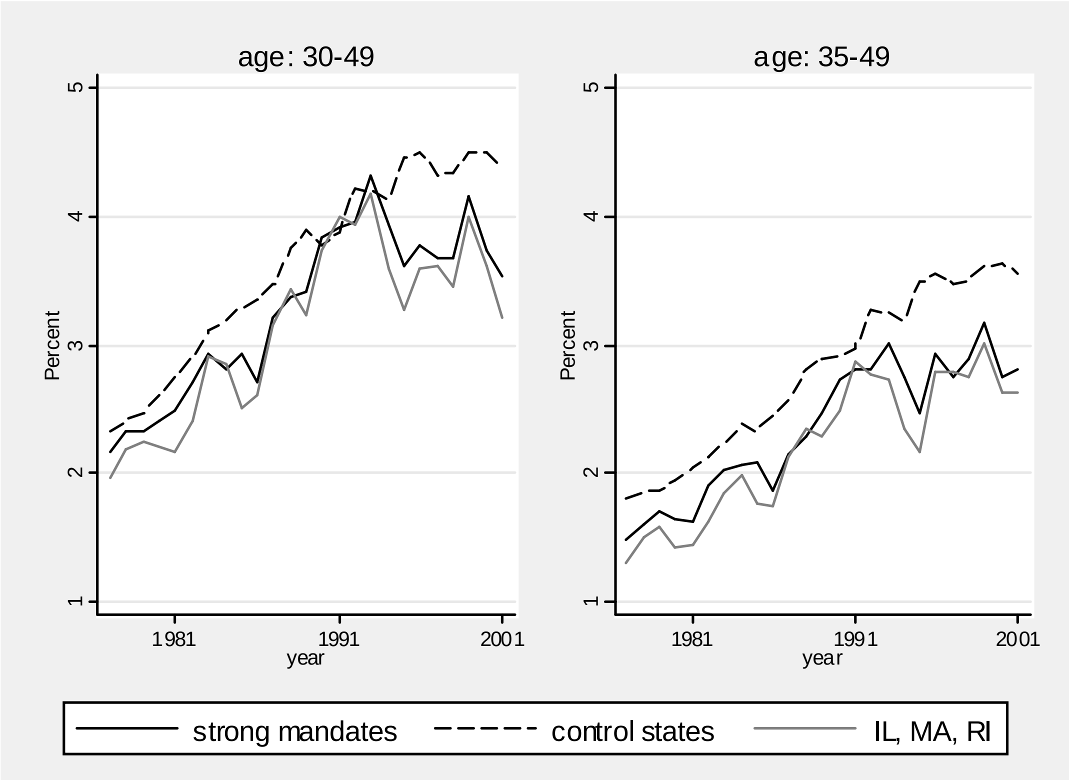

Figure 7: Percentage of women who are married without children

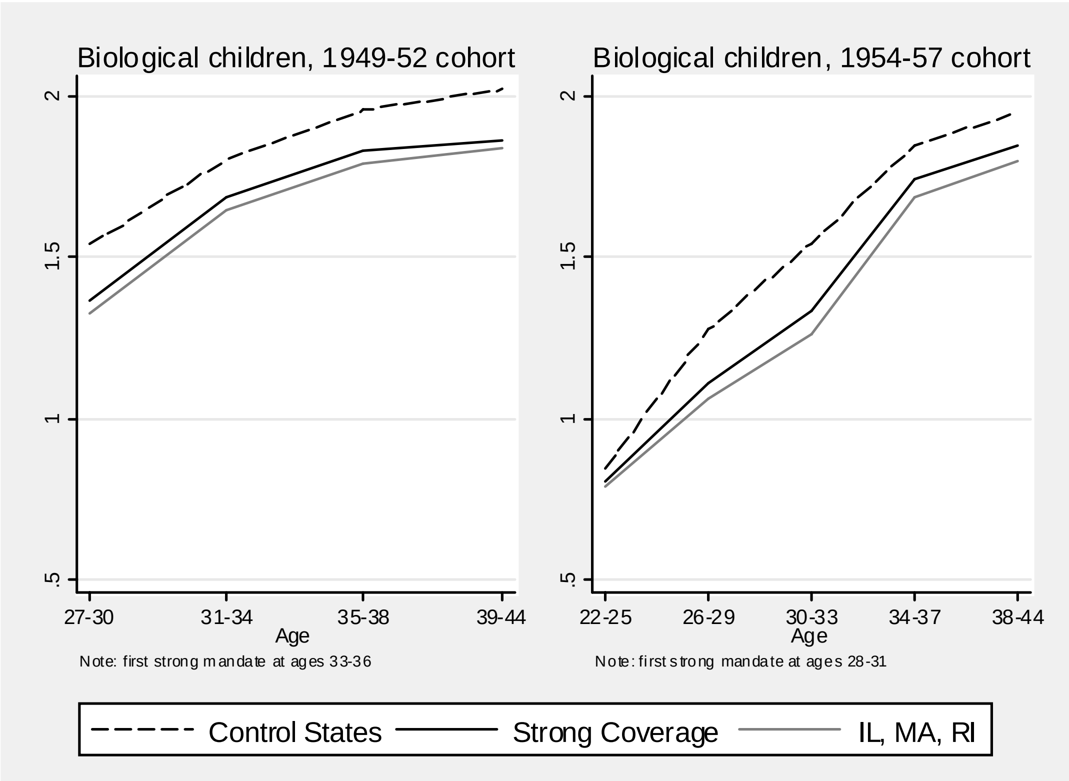

Figure 8: Average number of own biological children for two different cohorts

Table 2: The effect of infertility insurance coverage mandates on new mothers' age. OLS estimates.

	All States		**Strong Coverage and Control States**		**MA, IL, RI and Control States**	
	All	**Whites**	**All**	**Whites**	**All**	**Whites**
Mandated Coverage						
1-5 years	-0.032	-0.111	0.150	0.179	0.201	0.215
	(0.101)	(0.137)	(0.128)	(0.157)	(0.158)	(0.183)
6-10 years	0.013	-0.098	0.374	0.420	0.430	0.449
	(0.144)	(0.195)	(0.244)	(0.302)	(0.307)	(0.365)
More than 10 years	0.027	-0.135	0.672**	0.877**	1.217***	1.379***
	(0.235)	(0.301)	(0.336)	(0.391)	(0.184)	(0.198)
F-test of joint significance	[0.139]	[0.124]	[0.036]	[0.024]	[< 0.001]	[< 0.001]
t-tests of equality:						
"1-5" vs. "6-10" coeff.	[0.251]	[0.441]	[0.035]	[0.055]	[0.074]	[0.108]
"6-10" vs. "More than 10"	[0.458]	[0.591]	[0.048]	[0.023]	[< 0.001]	[< 0.001]
"1-5" vs. "More than 10"	[0.380]	[0.541]	[0.015]	[0.007]	[< 0.001]	[< 0.001]
N. Obs.	45,433,427	36,898,570	30,181,251	24,369,771	28,655,352	23,369,135

Note: All models include year and state fixed effects. Levels of statistical significance: *** denotes significance at the 1-percent level; ** at the 5-percent level; and * at the 10-percent level. Standard errors, displayed in round brackets, are clustered at the state level. P-values corresponding to the F-tests of joint significance and the one-sided t-tests of equality are displayed in square brackets. Data: Birth certificates from the National Vital Statistics (National Center for Health Statistics). Only one observation was kept per delivery. We uniquely identify multiple-birth mothers by using, whenever available, various variables such as year, month and day of birth, gestation time, state, county and place or facility of birth, attendant at birth, plurality, maternal age, race, years of schooling, marital status, place of birth, state, county, city and smsa of residence and paternal age and race.

Table 3: Estimated state weights in the synthetic control group

	Strong		**IL-MA-RI**	
	All	**Whites**	**All**	**Whites**
Alabama	0	0	**0.043**	0
Alaska	**0.11**	0	**0.02**	0
Arizona	**0.148**	**0.105**	0	0
Colorado	0	**0.038**	0	**0.111**
Delaware	0	0	0	0
District of Columbia	**0.074**	**0.013**	**0.045**	**0.011**
Florida	0	0	0	0
Georgia	0	0	0	0
Idaho	0	0	0	0
Indiana	0	**0.001**	0	0
Iowa	0	0	0	0
Kansas	0	0	0	0
Kentucky	0	0	0	0
Louisiana	0	0	**0.044**	**0.001**
Maine	0	0	0	0
Michigan	**0.052**	**0.111**	**0.007**	**0.057**
Minnesota	**0.106**	**0.168**	**0.105**	**0.133**
Mississippi	0	0	0	0
Missouri	0	0	0	0
Nebraska	0	0	0	0
Nevada	**0.02**	**0.006**	0	0
New Hampshire	0	0	0	0
New Jersey	**0.301**	**0.262**	**0.413**	**0.392**
New Mexico	0	0	0	0
North Carolina	**0.035**	**0.066**	0	0
North Dakota	**0.007**	0	0	0
Oklahoma	0	0	0	0
Oregon	0	0	0	0
Pennsylvania	0	0	0	0
South Carolina	**0.075**	0	0	0
South Dakota	0	0	0	0
Tennessee	0	0	0	0
Utah	**0.014**	0	**0.006**	0
Vermont	**0.037**	0	**0.151**	**0.093**
Virginia	0	**0.075**	0	0
Washington	0	**0.098**	0	0
Wisconsin	**0.021**	**0.002**	0	0
Wyoming	0	**0.055**	**0.166**	**0.202**

Table 4: Maternal age at first birth predictor means. Sample: all new mothers.

	States with Strong Coverage		**Average of**
	Real	Synthetic	**Control States**
mean % of new mothers age >35 (1981-83)	0.01396	0.01436	0.01514
mean % of new mothers age >35 (1977-80)	0.02131	0.02164	0.00981
mean % married women (1982-84)	0.52436	0.52577	0.55884
mean abortion rate (1978-1982)	29.1949	31.8491	24.4162
mean % white females (1982-84)	0.81651	0.81951	0.85703
mean % white females (1977-81)	0.82268	0.82580	0.84785
mean % black females (1981-84)	0.13906	0.14113	0.13269
mean % black females (1977-80)	0.13639	0.13620	0.12837
mean % highly educated women (1982-84)	0.36872	0.36441	0.31850
mean % highly educated women (1977-81)	0.31365	0.31367	0.28089
mean female employment rate (1982-84)	0.61664	0.61537	0.59490
mean female participation rate (1977-84)	0.64740	0.64819	0.63704
mean previous year female log hourly wage (1982-84)	1.95119	1.94975	1.84945
mean previous year female employment rate (1983-1984)	0.65491	0.65574	0.64222
mean previous year female employment rate (1977-1982)	0.63055	0.62794	0.62138
mean % of women covered by ESI in own name (1982-84)	0.34135	0.36550	0.33499
Maternal age at first birth, 1984	23.9426	23.9471	23.3413
Maternal age at first birth, 1982	23.5474	23.5310	22.9629
Maternal age at first birth, 1981	23.3371	23.3320	22.7918
Maternal age at first birth, 1979	22.9212	22.9489	22.4131
Maternal age at first birth, 1977	22.6122	22.6055	22.0579
Maternal age at first birth, 1976	22.4229	22.4288	21.8879
Maternal age at first birth, 1975	22.2059	22.2042	21.6589
Maternal age at first birth, 1974	22.0984	22.0898	21.5136
Maternal age at first birth, 1973	21.9157	21.9237	21.3490
Maternal age at first birth, 1972	21.8149	21.8108	21.2541

Notes: Each predictor variable is averaged for the period(s) indicated. ESI stands for employment sponsored health insurance.

Table 5: Maternal age at first birth predictor means. Sample: White new mothers in strong coverage and control states.

	States with Strong Coverage		**Average of**
	Real	Synthetic	**Control States**
mean % of new mothers age >35 (1981-83)	0.02245	0.02196	0.01595
mean % of new mothers age >35 (1977-80)	0.01435	0.01401	0.01017
mean % married women (1982-84)	0.56093	0.56378	0.59635
mean abortion rate (1978-1982)	29.0041	28.8335	24.0818
mean % highly educated women (1982-84)	0.38035	0.37813	0.33047
mean % highly educated women (1977-81)	0.32129	0.32216	0.29186
mean female employment rate (1982-84)	0.63634	0.63358	0.61246
mean female participation rate (1977-84)	0.65595	0.65609	0.64175
mean previous year female log hourly wage (1982-84)	1.94159	1.91633	1.85611
mean previous year female employment rate (1983-1984)	0.67154	0.67638	0.63228
mean previous year female employment rate (1977-1982)	0.64465	0.64003	0.65999
mean % of women covered by ESI in own name (1982-84)	0.34026	0.35837	0.33873
Maternal age at first birth, 1984	24.4470	24.4412	23.7488
Maternal age at first birth, 1982	24.0012	23.9802	23.3261
Maternal age at first birth, 1981	23.7702	23.7648	23.1451
Maternal age at first birth, 1979	23.3631	23.3934	22.7824
Maternal age at first birth, 1977	23.0335	23.0371	22.4308
Maternal age at first birth, 1976	22.8416	22.8643	22.2657
Maternal age at first birth, 1975	22.6107	22.6152	22.0325
Maternal age at first birth, 1974	22.6152	22.4881	21.8878
Maternal age at first birth, 1973	22.3086	22.3041	21.7171
Maternal age at first birth, 1972	22.1955	22.1814	21.6227

Notes: Each predictor variable is averaged for the period(s) indicated. ESI stands for employment sponsored health insurance.

Table 6: Maternal age at first birth predictor means. Sample: all new mothers in Illinois, Massachussetts and Rhode-Island and control states.

	IL-MA-RI		**Average of**
	Real	Synthetic	**Control States**
mean % of new mothers age >35 (1984)	0.03086	0.03021	0.02109
mean % of new mothers age >35 (1981-83)	0.02193	0.02171	0.01514
mean % of new mothers age >35 (1977-80)	0.01428	0.01432	0.00981
mean % married women (1982-86)	0.51522	0.53808	0.55649
mean abortion rate (1985)	26.7785	30.5249	22.7604
mean abortion rate (1978-1982)	28.2173	30.8581	24.4162
mean % white females (1982-86)	0.86303	0.86192	0.84451
mean % white females (1977-81)	0.86897	0.86984	0.85703
mean % black females (1981-86)	0.11699	0.11586	0.13507
mean % black females (1977-80)	0.11656	0.11421	0.12837
mean % highly educated women (1982-86)	0.38548	0.37434	0.32935
mean % highly educated women (1977-81)	0.32112	0.31028	0.28089
mean female employment rate (1982-86)	0.62411	0.62739	0.61164
mean female participation rate (1977-86)	0.65378	0.65402	0.64923
mean previous year female log hourly wage (1982-86)	1.9786	1.92351	1.86348
mean previous year female employment rate (1983-1986)	0.66094	0.66860	0.66012
mean previous year female employment rate (1977-1982)	0.63202	0.62750	0.62138
mean % of women covered by ESI in own name (1982-86)	0.34576	0.34579	0.34118
Maternal age at first birth, 1986	24.5725	24.5688	23.6388
Maternal age at first birth, 1984	24.1552	24.1534	23.3413
Maternal age at first birth, 1982	23.7240	23.6820	22.9629
Maternal age at first birth, 1981	23.4829	23.4686	22.7918
Maternal age at first birth, 1979	23.0364	23.0830	22.4131
Maternal age at first birth, 1977	22.7486	22.7613	22.0579
Maternal age at first birth, 1976	22.5770	22.6002	21.8879
Maternal age at first birth, 1975	22.3530	22.3657	21.6589
Maternal age at first birth, 1974	22.3657	22.2320	21.5136
Maternal age at first birth, 1973	22.0657	22.0522	21.3490
Maternal age at first birth, 1972	21.9395	21.9356	21.2541

Notes: Each predictor variable is averaged for the period(s) indicated. ESI stands for employment sponsored health insurance.

Table 7: Maternal age at first birth predictor means. Sample: White only new mothers in Illinois, Massachussetts and Rhode-Island.

	IL-MA-RI		**Average of**
	Real	Synthetic	**Control States**
mean % of new mothers age >35 (1984)	0.03306	0.03147	0.02238
mean % of new mothers age >35 (1981-83)	0.02319	0.02252	0.01595
mean % of new mothers age >35 (1977-80)	0.01475	0.01454	0.01017
mean % married women (1982-86)	0.54643	0.56190	0.59451
mean abortion rate (1985)	26.9639	29.033	22.4556
mean abortion rate (1978-1982)	28.3668	29.0851	24.0818
mean % highly educated women (1982-86)	0.39357	0.39674	0.34201
mean % highly educated women (1977-81)	0.32586	0.32592	0.29186
mean female employment rate (1982-86)	0.64702	0.64609	0.62856
mean female participation rate (1977-86)	0.66582	0.66575	0.65437
mean previous year female log hourly wage (1982-86)	1.97195	1.92088	1.8710
mean previous year female employment rate (1983-1986)	0.68238	0.69066	0.67701
mean previous year female employment rate (1977-1982)	0.64789	0.64421	0.63228
mean % of women covered by ESI in own name (1982-86)	0.34774	0.34508	0.34502
Maternal age at first birth, 1986	25.0893	25.0598	24.0802
Maternal age at first birth, 1984	24.6304	24.6250	23.7488
Maternal age at first birth, 1982	24.1502	24.1109	23.3261
Maternal age at first birth, 1981	23.8924	23.8705	23.1451
Maternal age at first birth, 1979	23.4595	23.5148	22.7824
Maternal age at first birth, 1977	23.1485	23.1591	22.4308
Maternal age at first birth, 1976	22.9762	23.0147	22.2657
Maternal age at first birth, 1975	22.7394	22.7539	22.0325
Maternal age at first birth, 1974	22.6131	22.6100	21.8878
Maternal age at first birth, 1973	22.4344	22.4189	21.7171
Maternal age at first birth, 1972	22.3136	22.2913	21.6227

Notes: Each predictor variable is averaged for the period(s) indicated. ESI stands for employment sponsored health insurance.

Table 8: The long run impact of strong infertility insurance coverage mandates on the age of new mothers

	Raw	**Synthetic control group estimates**				
	DID (2001)	Parameter estimate (2001)	p-value	p-value5	RMSPE	p-value of ratio
		Panel A: Strong Mandates				
All	0.42	0.266	0.158	0.094*	0.0177	0.026**
Whites	0.43	0.317	0.211	0.161	0.0179	0.053*
		Panel B: Illinois, Massachusetts and Rhode Island				
All	0.43	0.341	0.105	0.091*	0.0237	0.079*
Whites	0.42	0.448	0.158	0.114	0.0269	0.053*

Notes: Treatment is assumed to start in 1985 for states with strong mandates and in 1987 for IL, MA and RI. Levels of statistical significance: *** denotes significance at the 1-percent level; ** at the 5-percent level; and * at the 10-percent level. RMSPE denotes the root mean squared prediction error. Raw DID refers to difference-in-differences estimates obtained using the same aggregate data and not controlling for any additional variables. All the p-values displayed are based on placebo runs that are described in Section 4.2. Predictors used in estimation of the synthetic control effect are described in tables Tables 4, 5, 6 and 7.

Table 9: Evolution of the synthetic control gap in maternal age at first birth. OLS estimates.

	Strong Mandates		**IL, MA, RI**	
	All	**Whites**	**All**	**Whites**
Mandated Coverage:				
1-5 years	0.093***	0.114***	0.048	0.095***
	(0.016)	(0.016)	(0.029)	(0.027)
6-10 years	0.158***	0.158***	0.100**	0.172***
	(0.007)	(0.010)	(0.041)	(0.045)
More than 10 years	0.261***	0.279***	0.292***	0.378***
	(0.012)	(0.010)	(0.017)	(0.024)
F-test of joint significance	[< 0.001]	[< 0.001]	[< 0.001]	[< 0.001]
t-tests of equality:				
"1-5" vs. "6-10" coeff.	[0.002]	[0.036]	[0.324]	[0.163]
"6-10" vs. "More than 10" coeff.	[< 0.001]	[< 0.001]	[0.0011]	[0.0017]
"1-5" vs. "More than 10" coeff.	[< 0.001]	[< 0.001]	[< 0.001]	[< 0.001]
N. Obs.	17	17	15	15
R^2	0.982	0.980	0.881	0.924

Note: The dependent variable is the post-treatment maternal age at first birth gap between states with strong coverage vs. the synthetic control group.Levels of statistical significance: *** denotes significance at the 1-percent level; ** at the 5-percent level; and * at the 10-percent level. Robust standard errors are displayed in round brackets. P-values corresponding to the F-tests of joint significance and the one-sided t-tests of equality are displayed in square brackets. The model includes no constant.

Table 10: The effect of strong infertility insurance coverage mandates on the probability of having at least one child by age 30/35. Probit marginal effects

	Prob. at least one child by 30				**Prob. at least one child by 35**			
	All	**White**	**All**	**White**	**All**	**White**	**All**	**White**
Panel A: Mandated Coverage at age 30/35:								
1-5 years	0.0040	0.0144*			−0.0230*	−0.0198		
	(0.0071)	(0.0081)			(0.0120)	(0.0137)		
6-10 years	−0.0290***	−0.0188*			−0.0127	−0.0033		
	(0.0098)	(0.0116)			(0.0100)	(0.0149)		
Panel B: Mandated Coverage at the interview								
1-5 years			−0.0092	−0.0025			−0.0099	−0.0274
			(0.0101)	(0.0112)			(0.0109)	(0.0663)
6-10 years			−0.0358***	−0.0240*			−0.0310***	−0.0732
			(0.0101)	(0.0126)			(0.0111)	(0.0482)
N. of Obs	109,211	93,121	109,211	93,121	67,618	57,953	67,618	57,953
Pseudo R^2	0.118	0.128	0.118	0.128	0.117	0.131	0.117	0.131
Log-Lik.	−50,006	−42,699	−50,001	−42,698	−24,769	−21,091	−24,769	−21,092

Notes: Treatment is assumed to start in 1985 for all treated states. Levels of statistical significance: *** denotes significance at the 1-percent level; ** at the 5-percent level; and * at the 10-percent level. All regressions control for year dummies, state fixed effects, education variables (high-school, more than high-school), working status, not married status, and age dummies. Regressions for all women also include race dummies.

Table 11: The effect of IL, MA and RA infertility insurance coverage mandates on the probability of having at least one child by age 30/35. Probit marginal effects

	Prob. at least one child by 30				Prob. at least one child by 35			
	All	**White**	**All**	**White**	**All**	**White**	**All**	**White**
Panel A: Mandated Coverage at age 30/35:								
1-5 years	0.0115***	0.0210***			−0.0199	−0.0164		
	(0.0044)	(0.0050)			(0.0176)	(0.0171)		
6-10 years	−0.0335***	−0.0223***			0.0043	0.0160		
	(0.0074)	(0.0082)			(0.0151)	(0.0102)		
Panel B: Mandated Coverage at the interview								
1-5 years			−0.0048	0.0017			−0.0096	−0.0037
			(0.0099)	(0.0116)			(0.0121)	(0.0155)
6-10 years			−0.0333***	−0.0243			−0.0199	−0.0062
			(0.0143)	(0.0167)			(0.0203)	(0.0176)
N. of Obs	103,499	89,337	103,499	89337	63,971	55,498	63,971	55,498
Pseudo R^2	0.118	0.1278	0.118	0.128	0.119	0.131	0.118	0.131
Log-Lik.	−47,396	−40,997	−47,395	−40,997	−23,465	−20,247	−23,466	−20,248

Notes: Treatment is assumed to start in 1987 for IL, MA and RI. Levels of statistical significance: *** denotes significance at the 1-percent level; ** at the 5-percent level; and * at the 10-percent level. All regressions control for year dummies, state fixed effects, education variables (high-school, more than high-school), working status, not married status, and age dummies. Regressions for all women also include race dummies.

Table 12: The impact of the mandates on population structures. Raw DID estimates.

	Percentage of Women							
	All				Married and without children			
	30-49 years old		35-49 years old		30-49 years old		35-49 years old	
	DID	Percent	DID	Percent	DID	Percent	DID	Percent
Arkansas	0.003	1.05	−0.02	−11.36	0.001	3.78	−0.002	−8.91
Hawaii	−0.015	−4.91	−0.011	−5.67	0.016	52.76	0.016	70.56
Illinois	−0.01	−3.33	−0.01	−4.79	−0.018	−40.86	−0.013	−40.11
Maryland	−0.044	−14.32	−0.049	−20.81	−0.022	−49.45	−0.022	−69.00
Massachusetts	0.025	9.15	0.011	5.91	−0.004	−14.73	0	1.21
Rhode-Island	0.017	6.10	0.014	7.05	0.004	14.06	0.008	64.74
Strong states	−0.001	−0.41	−0.019	−4.68	−0.005	−17.07	−0.005	−24.07
IL, MA, RI	0.002	0.702	0.005	2.48	−0.009	−27.32	−0.005	−22.95

Notes: Columns labeled "DID" show unconditional difference-in-differences estimates where the evolution of the variable between 2001 and the treatment year for each state is compared with the evolution of the variable for all control states during the same period. Columns labeled "Percent" display the percentage that the DID estimate represents in terms of the level of the variable in the treatment year. Treatment is assumed to start in 1985 for the combination of all states in the row "Strong states" and in 1987 for the group of states labeled "Comprehensive states" (i.e. IL, MA, RI)

Table 13: The effect of infertility insurance coverage mandates on the number of biological children, Zero nflated poisson marginal effects. Women aged 44 only.

	Strong Coverage and Control States		**MA, IL, RI and Control States**	
	All	**Whites**	**All**	**Whites**
Mandated Coverage:				
1-5 years	0.134	0.121	0.114	0.124
	(0.144)	(0.119)	(0.181)	(0.141)
6-10 years	−0.076	−0.078	−0.068	−0.119
	(0.123)	(0.138)	(0.166)	(0.165)
More than 10 years	−0.047	−0.038	−0.076	−0.094
	(0.150)	(0.146)	(0.101)	(0.086)
% of zeros	13.11	13.10	13.21	13.14
Vuong test p-value	[< 0.001]	[< 0.001]	[< 0.001]	[< 0.001]
Log-likelihood	−15,322.0	−12,870.6	−14,556.6	−12,385.4
N. Obs.	8,609	7,365	8,163	7,068

Note: All models include year and state fixed effects as well as educational attainment indicators, a binary variable indicating whether the woman works or not and a non-married dummy variable. Regressions for all women include race dummies as well. Levels of statistical significance: *** denotes significance at the 1-percent level, ** at the 5-percent level, and * at the 10-percent level. Standard errors, displayed in round brackets, are clustered at the state level.

Table 14: The effect of infertility insurance coverage mandates on the number of biological children, Zero inflated poisson marginal effects. All women.

	Strong Coverage and Control States		**MA, IL, RI and Control States**	
	All	**Whites**	**All**	**Whites**
Mandated Coverage:				
1-5 years	-0.009	0.000	-0.017	-0.007
	(0.021)	(0.032)	(0.025)	(0.038)
6-10 years	-0.048^{**}	-0.042	-0.051	-0.055
	(0.023)	(0.036)	(0.032)	(0.046)
More than 10 years	-0.037	-0.031	0.001	-0.012
	(0.027)	(0.022)	(0.045)	(0.033)
% of zeros	36.15	37.22	36.17	37.26
Vuong test p-value	[< 0.001]	[< 0.001]	[< 0.001]	[< 0.001]
Log-likelihood	$-379,086.3$	$-309,272.6$	$-359,924.6$	$-297,449.2$
N. Obs.	288,770	242,707	274,018	233,260

Note: All models include year and state fixed effects as well as age, educational attainment indicators, a binary variable indicating whether the woman works or not and a non-married dummy variable. Regressions for all women include race dummies as well. Levels of statistical significance: *** denotes significance at the 1-percent level, ** at the 5-percent level, and * at the 10-percent level. Standard errors, displayed in round brackets, are clustered at the state level.

References

[1] Abadie, Alberto and Javier Gardeazabal (2003): "The Economic Costs of Conflict: A Case Study of the Basque Country", American Economic Review, 93(1), 112-132.

[2] Abadie, Alberto, Diamond, Alexis, and Jens Hainmueller (2010): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," Journal of the American Statistical Association, 105, 493-505.

[3] Benyamini, Yael (2003): "Hope and Fantasy among Women Coping with Infertility and its Treatments," In *Between Stress and Hope: from a disease centered to a health-centered perspective* edited by Rebecca Jacoby and Giora Keinan. Praeger series in health psychology . Praeger Publishers, Westport, CT, USA.

[4] Bhattacharya, Jay, and William Vogt (2000): "Could we tell if health insurance mandates cause unemployment? A note on the literature," Mimeo.

[5] Bitler, Marianne (2005): "Effects of Increased Access to Infertility Treatment to Infant and Child Health Outcomes: Evidence from Health Insurance Mandates," RAND Labor and Population, WR-330.

[6] Bitler, Marianne and Lucie Schmidt (2011): "Utilization of Infertility Treatments: The Effects of Insurance Mandates," Conditionally accepted at Demography.

[7] Blau, F.D.,and L.W. Kahn (1997): "Swimming Upstream: Trends in the Gender Wage Differential in the 1980s," Journal of Labor Economics 15, 1–43.

[8] Buckles, Kasey (2005): "Stopping the Biological Clock: Infertility Treatments and the Career-Family Tradeoff," BU Dissertation.

[9] Bundorf, M. Kate, Henne, Melinda, and Laurence Baker (2007): "Mandated Health Insurance Benefits and the utilization and Outcomes of Infertility Treatments," NBER Working Paper 12820.

[10] 1996 Assisted Reproductive Technology Success Rates, National Summary and Fertility Clinic Reports. Centers for Disease Control and Prevention, Division of Reproductive Health, Atlanta Georgia.

[11] 2005 Assisted Reproductive Technology Success Rates, National Summary and Fertility Clinic Reports. Centers for Disease Control and Prevention, Division of Reproductive Health, Atlanta Georgia.

[12] 2006 Assisted Reproductive Technology Success Rates, National Summary and Fertility Clinic Reports. Centers for Disease Control and Prevention, Division of Reproductive Health, Atlanta Georgia.

[13] 2007 Assisted Reproductive Technology Success Rates, National Summary and Fertility Clinic Reports. Centers for Disease Control and Prevention, Division of Reproductive Health, Atlanta Georgia.

[14] Caucutt, Elizabeth M., Nezih Guner, and John Knowles (2002): "Why Do Women Wait? Matching, Wage Inequality and the Incentives for Fertility Delay," Review of Economic Dynamics, 5, pp 815-855.

[15] Collins, John (2001): "Cost-Effectiveness of in vitro fertilization," Seminars in Reproductive Medicine, 19(3), 279-289.

[16] Friedberg, Leora (1998): "Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data," American Economic Review, 88, 608-627.

[17] Goldin C. and L. F. Katz (2002): "The Power of Pill: Oral Contraceptives and Women's Career and Marriage Decisions," Journal of Political Economy, vol 110(4), pp 730-770.

[18] Grant, Jonathan with Stijn Hoorens, Federico Gallo and Jonathan Cave (2006): Should Art be Part of a Policy Mix", Project Resource, RAND Europe.

[19] Gruber, Jonathan (2004): "Is making Divorce Easier Bad for Children?". Journal of Labor Economics, 22 (4), 799-833.

[20] Hamilton, Barton H. and Brian McManus (2005):"Infertility Treatment Markets: The Effects of Competition and Policy," Mimeo.

[21] Hao, Lingxin and Andrew J. Cherlin (2004): "Welfare Reform and Teenage Pregnancy, Childbirth, and School Dropout," Journal of Marriage and Family, 66, 179-194.

[22] Heckman, J., Ichimura, H., Smith, J., and Petra Todd (1998): "Characterizing Selection Bias Using Experimental Data," Econometrica 66(5), 1017-1098.

[23] Heckman, J., Ichimura, H., and Petra Todd (1997): "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," The Review of Economic Studies 64(4), 605-654.

[24] King, Miriam, Ruggles, Steven, Alexander, J. Trent, Flood, Sarah, genadek, Katie, Schroeder, Matthew B., Trampe, Brandon and Rebecca Vick. 2010. Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota.

[25] International Federation of Infertility Societies (IFFS) Surveillance 07, edited by Jones, Howard W. JR., Cohen, Jean, Cooke, Ian and Kempers, Roger. Supplement to Fertility and Sterility, April, 2007, vol 87, sup. 1.

[26] Lampi, Elena (2006): "The personal and general risks of age-related female infertility: Is there an optimistic bias or not?," Working Paper in Economics No. 231, School of Business, Economics and Law, Göteborg University.

[27] Mead, Lawrence M. (2004): "State Political Culture and Welfare Reform," The Policy Studies Journal, Vol. 32, No. 2, 271-296.

[28] Michalopoulos, Charles, Bloom, H., and C. Hill (2004): "Can Propensity Score Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?," Review of Economics and Statistics, vol86(1), pp 156-179.

[29] Mookim, Pooja G., Ellis, Randall P. and Ariella Kahn-Lang (2008): "Infertility Treatment, ART and IUI Procedures and Delivery Outcomes: How Important is Selection?," Mimeo, Boston University.

[30] Neumann, P. J., Gharib, S. D. and Weinstein, M. C. (1994): "The Cost of a Successful delivery with in vitro fertilization," *The New England Journal of Medicine*, 331(4), 239-43.

[31] Ohinata, Asako (2009): "Did the US infertility insurance mandates affect the time of first birth?," University of Warwick, mimeo.

[32] Olivetti, Claudia (2006): "Changes in Women's Aggregate Hours of Work: The Role of Returns to Experience," Review of Economic Dynamics 9, no. 4, 557-587.

[33] Schmidt, L. (2005): "Infertility Insurance Mandates and Fertility," *Journal of Health Economics* 26, pp 431-446.

[34] Schmidt, L. (2007): "Effects of Infertility Insurance Mandates on Fertility," American Economic Review 95(2), pp 204-208.

[35] Schmidt, L. (2005b): "Effects of Infertility Insurance Mandates on Fertility," Labor and Demography 0511014 WPA.

[36] Simon, Kosali Ilayperuma (2004): "Adverse selection in health insurance markets? Evidence from state small-group health insurance reforms," Journal of Public Economics, 89(9-10), pp 1865-1877.

[37] Smith, Jeffrey and Petra Todd (2005): "Does Macthing Overcome Lalonde's Critique of Nonexperimental Estimations," Journal of Econometrics March-April, pp 305-353.

[38] Ziebe, Søren, and Paul Devroey (2008): "Assisted Reproductive Tecnhologies are an Integrated Part of National Strategies Addressing Demographic and Reproductive Challenges," in Human Reproduction Update, pp1-10, European Society of Human Reproduction and Embriology, Oxford University Press.

9 Appendix - Changes in Infertility Treatment Laws

Four out of the six strongly treated states (Arkansas, Hawaii, Maryland and Massachusetts) revised their mandates during our sample period, *i.e.*, before 2001. Table 15 briefly describes these revisions. The revisions in Arkansas and Maryland reduce coverage and hence would tend to decrease the estimated impact of the original mandates. Massachusetts' revision in 1995 established that the IVF procedures ICSI and ZIFT should be covered. ICSI is a particularly effective IVF procedure in cases of male infertility in which a single sperm is injected directly into an egg. Because this procedure was invented in 1991, it could not have been explicitly contemplated in the original mandate, although Massachusetts's original mandate covered IVF procedures. ICSI accounts for a large percentage of the fresh non-donor eggs or embryos, 57.9% according to CDC, 2001. The usage of ZIFT, however, has been declining gradually and in 2001 it accounted for less than 2% of ART procedures (CDC, 2001). Finally, the Hawaiian revision in 1995 is clearly an expansion of coverage to dependant non-married individuals. Simple DID estimates of the effect of this revision on the marriage probability of first-time mothers show either no effect or a positive effect for Whites.[20] Hence, since there is no evidence that the Hawaiian revision decreased the marriage rate in the state, it is unlikely that it had a significant impact on the number of covered users.

[20] These DID estimates were obtained using the natality files data from 1987 (the date of the original mandate) to 2001, taking Hawaii as the treatment group against all non-treated states as controls. When controlling for race and education, we obtain essentially the same effect. We repeated the estimation for Whites only and obtained a positive and significant effect of the 1995 revision on marriage rates of first-time mothers, which goes in the opposite direction of what would be expected.

State	Original Mandate	Revision	Description of revisions
Arkansas	1987	1991	Imposition of minimum and maximum benefits and setting standards
Hawaii	1987	1995	Patient does not have to be the spouse of the insured but only a dependen
Maryland	1985	1994	Exempt businesses $\leq$ 50 employees from IVF coverage
		2000	Restricts coverage to 3 IVF attempts/live birth
			Exempts organizations with religious conflicts
Mass.	1987	1995	Extends coverage to ICSI and ZIFT*

Notes: Sources: Schmidt b); https://www.hrtools.com/
http://us.firstvisitivf.org/display.asp?page=IVF_coverage_in_USA#state-law
http://www.resolve.org/
http://www.fertilitylifelines.com/payingfortreatment/state-mandatedinsurancelist.jsp
*For Massachusetts, see 1995-08 211 CMR 37.00, New Infertility Mandated Benefits

Table 15: Revisions of infertility treatment mandates